A)

	Motif 1	Motif 2	Motif 3	Motif 4
Rv2837c_Mt	[35] VGVVCHVHPEAETIGAGLALALVLDGCG	VDLVVTVDEIPSVDRLGALG	[6] RELLVIDHHASND	[6] SADSTTTMVAEILDAWGKPIDPRVAHCIYAGLATDTGSFRWASV
sll1253_Ssp	[34] DLILCHQTADFEVLGAAVGLAKLHPGSR	IRSLYIVDENQQGDRLGKAA	[8] RQVAIYIVDHHLNSP	[11] AVGASTTLIVEKLQRADISLSMVEASVMALGIHVDTGSLIFTQT
MGPA_Mg	[32] IVIFHHVRPGEVCLGAQQGLFHLIKANF	EALAIVDEANYKNRIELRE	[7] KAVLRIDHHPNED	[11] SYVACCEQIVEMATVAKWTIPPVAATLLYIGIYVDSNRFLYSNT
YTQI_Bs	[30] IILRHVRPEPDAVGSQCGLTEILRETY	GALVIVCDTANQERIDDQR	[4] AKLMKIDHHPNED	[13] SVSEMIYELYLEGKEHGWKLNTKAAELIYAGIVGDTGRFLFPNT
AF2029_Af	[37] LGIFTHDNPDPDSMSSAYALREIAKQFD	YDAFAIVDSSGPGVNNSIP	[3] DISIVIDHHPAEK	[10] DVGATAILTEYIKELKITPSKILATALFFGIKSETDEFKRNTR
MJ0977_Mj	[23] NKLIVTHIDTEGLTSRAILQKLAERLN	YDLIIPADLGSGQLRMIKE	[11] DKIIILDHHQPEE	[18] GAEICGAGVSYLFAKAINNDWIDLAKYAVLGAVGDIQNIEGKLI
MJ1198_Mj	[46] RPIIIRHAPETGYCGLELEKAILPII	LPLVLLDLNGSTDEDIPAI	[7] IEVIVIDHHFPGE	[85] KGRTYDREYLEKIALCMDPEAFYLRFMDGKGIVDEILATNIKEF
HP1042_Hp	[27] MQVYHLSHIDLDGYACQLVSKQFFKNIQ	EFLIIVSDLNLNLNEAEYL	[13] IQIQLIDHHISGK	[19] IVYEFLKKHYAILEPKNTTWLEPLVEMVNSVDIWDTQGYGFELG
RecJ_Hi	[31] QKIVIVGDHDADGATSTALSVLALRQLG	VQLLMTVDNGVSSFDGVAF	[5] IRVLVTIDHHLPPE	[33] LAVRAKFRELGIFTAETQPNFTDLLDLVALGTIADVVPLDQNNR
RecJ_Hp	[27] TEILVVGDYDADGVISSAIMAKFFESLN	APLIITVDNGINAFEAARF	[5] YTLIITIDHHCLHH	[27] LVAFYLCYGIHQLLGKEKSHSSELLCLAGVATIPDMMPLTFFNR
RecJ_Ssp	[32] EKVEIWGDPEAEGITSTAVLMEGLGQFF	TKLIVTCDTGSTNLDEIVY	[5] MDVIVTIDHHTLPD	[27] VAFKLVEALYNQYPTVPQQPLEDLLDIVAIGELADLVTLQGDCR
YYBQ_Bs	[36] ILLIFGHQNPDTETCSAIAYADLKNKLG	VNGVILVDHNERQQSIKDI	[3] QVLEVIDHHRIAN	[12] PVGCTATILNKMYKENNVKIEKEIAGLMLSAIISDSLLFKSPTC
ICRA_Sg	[38] ILVFGHQHPDSEAIGSSYAFAYLAREAY	AEQVLTDHNEFQQSVADI	[3] EVVGVVDHHRVAN	[12] PVGSASSIVYRMFKEHSVAVSKEIAGLMLSGLISDTLLLKSPTT
MJ0608_Mj	[31] RYVVGHKNPETDSIASAIVLAYFLDCYP	GKEIILVDHSEKSQSFDDL	[3] KLIAIFIDHHKVGL	[19] IAELYFKDAIDLIGGKKKELKPDLAGLLLSAIISDTVLFKSPTT
AF0756_Af	[46] VTVVGHKNPETDSVCSAIAFAYLMNKWK	GKKVALVDHSEKAQTVDGI	[3] EVVAIVDHHKIGD	[12] PVGCTATVIKLLFDKTGVEIPKDIAGILLSSILSDTVIFKSATT
U60409_Lm	[44] TVVQGNEGQPMDSIVGCIYLAMLFDKQP	QIAHNIVDIAALNASVVLY	[14] RVVGVVDHHFDEQ	[11] LRTVGSACTLVTELYRECGEDVVCPTLTAPIVLDTVNFEPAQK
PPX1_Sc	[62] TICVGNESADVQSIASAITYSYCQYIYN	ELNSYLVDNNDTPKNLKNY	[2] NRVGLIDHHFDLQ	[14] SCSSLVFNYWYERLQGDREVVMNIAPLLMGALLIDTSNMRRKVE
PRUNE_Dm	[48] HLVMGNESOLEDSAVSAVTLAFVYAQRH	DVNVILVDHVSPLAPNVT	[1] NVTEILDHHPLED	[18] SVGSCATLVAQRYLAEDQPRSTSVAQLLHATIVLDTINFAPAAK
h-prune.1_Hs	[52] HVVLGNEACLEDSTVSALALAFYLAKT	QSTLILVDHHILSKSDTAL	[2] AVAEVLDHHPIEP	[12] LVGSCATLVTERILQGAPEILDRTAALLHGTILLDCVMMDLKIG

Family 1 / Family 2

	Specific motifs	GIs	
Rv2837c_Mt	[107] TVNLAAVASGFGGGGHRLAAGYTTTGS	1648883	
sll1253_Ssp	[123] DTDLTQLLEPYGGGGHAQAAAVNLRDV	1653244	
MGPA_Mg	[102] GINVRDIAIKYGGGGHNNASGAIITNK	1045875	
YTQI_Bs	[103] GPVINGLARKYNGGGGHPLASGASIYSW	2293259	
AF2029_Af	[102] EVLRRAFGDVGSAGGGHKFACGAYIPDN	2648507	
MJ0977_Mj	[254] AIKYASEKVNGSGGGHKFACGAYIPDN	2128614	
MJ1198_Mj	[109] QLMEEIPEASLDGGGHECAGSLKFVEG	2128720	
HP1042_Hp	[137] CDVCELSQMCFNGGGHRNASGGKIDGF	2314198	
RecJ_Hi	[168] RIHSQHPNMILKFGGGHAMAAGLSIREE	1172895	
RecJ_Hp	[155] DALNGVSSLLLGTYGGHRQACGLSVEKN	2313437	
RecJ_Ssp	[165] ALLHSQRHLMLGFGGHPFAAGLSLPLD	1652638	
YYBQ_Bs	[31] DLSKKTVEELISLDAKEPTLG	586817	[75] TALLKGVVSRKKQVVPVLTD
ICRA_Sg	[31] NLASKSAEEEIDIDARTPELN	1743856	[75] HAPLAGAVSRKKQVVPQLTE
MJ0608_Mj	[32] VVGKLKPEEIINMDFKNFDFN	1591318	[74] SVFLEGVMSRKKQVVPPLER
AF0756_Af	[32] AVDDLTAMDIIKRDYXDFDMS	11498362	[75] SVWLDGVMSRKKQVVPPLEK
U60409_Lm	[36] DVLALSVPQILRRDYXQFSFK	1407725	[105] YSLSDPSISRKK-LVPALSE
PPX1_Sc	[41] DIKGFSVSDILKKDYXQFNFQ	730369	[97] MFKQLNVEATRKQVVPYLEE
PRUNE_Dm	[37] DISKLLTEVIRKDMKVLQTD	1079081	[97] LRQHNVQATRKH-ILPIVKR
h-prune.1_Hs	[34] DVSGLTTBQMLRKDQKTIYRQ	24308263	[94] YLQGNTQVSRKK-LLPLLQE

Family 1 / Family 2

Fig. 1